Journal of the Korean
Statistical Society
Vol. 25, No. 4, 1996

An Optimum Design for First Order Fits to Correlated Responses†

Whasoo Bae [1]

Abstract

The aim of this paper is to find a suitable design which minimizes the expected discrepancy in fitting a first order model fearing quadratic terms as bias where there are more than two correlated responses. Kim and Draper(1994) discussed about choosing a design for straight line fits to two correlated responses. The general case with r responses is examined here and the result is applied to a specific case to help understandings.

Key Words : Multiple responses; Correlation; Design moment; Variance error; Bias error.

1. INTRODUCTION

Assume that an experimenter is to explore a functional relationship between predictors and r responses. And a first order model is to be fitted to r responses over the region of interest R, assumed to be a spherical region of

†This paper was supported by NON DIRECTED RESEARCH FUND, Korean Research Foundation, 1994.

[1]Department of Statistics, Inje University, Kimhae, Kyongnam, 621-749, Korea.

radius 1 here, where the true or feared relationship is a quadratic one over the whole region of operability, O. Then it is expected that there would be a certain amount of quadratic bias explaining the gap between the fitted model and the feared one. By choosing the suitable design which minimizes this gap, the first order model represents the true relationship satisfactorily within the region of interest so that it can be used for further purposes.

Box and Draper(1959) measured the gap between the fitted model and the feared one by using a criterion which devides the discrepancy into two parts , the bias error part and the variance error one. They discussed about choosing a design which minimizes these two types of errors, when the first order model is fitted to a response fearing the quadratic bias. In Box and Draper(1963), the selection of a second order rotatable design was discussed in single-response case. As seen in Khuri(1988), this type of problem has not been examined that much since 1963. Kim and Draper(1994) discussed about choosing a design for straight line fits to two correlated responses, considering the correlations between two responses. In this work, a suitable design is to be found in fitting a first order model to r responses fearing quadratic terms as bias.

Suppose that the model to be fitted is the first order polynomial with no common parameters so that that the estimation can be done using ordinary least squares estimation method.(See Box and Tiao(1973).) Then the model form is

$$\mathbf{Y} \;=\; \mathbf{X}\beta \;+\; \epsilon \;, \qquad (1.1)$$

where $\mathbf{Y} = (\mathbf{y_1}, \cdots, \mathbf{y_r})$ with $\mathbf{y}_i = (y_{1i}, \cdots, y_{Ni})'$, $\mathbf{X} = (\mathbf{x}_1', \mathbf{x}_2', \cdots, \mathbf{x}_N')'$ with $\mathbf{x}_u' = (1, x_{u1}, \cdots, x_{uk})'$, and where $\beta = (\beta_1, \cdots, \beta_r)$ with $\beta_i = (\beta_{0,i}, \beta_{1,i}, \cdots, \beta_{k,i})'$ for $i = 1, \cdots, r$ and $u = 1, \cdots, N$, where $\epsilon = (\epsilon_1, \cdots, \epsilon_r)$ with $\epsilon_i = (\epsilon_{1i}, \cdots, \epsilon_{Ni})'$ is an $N \times r$ error matrix and each row vector of ϵ has the variance-covariance matrix as Σ, where $\Sigma = (\rho_{ij}\sigma_i\sigma_j, i, j = 1, \cdots, r)$ with $\rho_{ij} = 1$ for $i = j$ and $\mathrm{cov}(\epsilon_{ui}, \epsilon_{vj}) = \rho_{ij}\sigma_i\sigma_j$ if $u = v$, 0 otherwise for any $i, j = 1, \cdots, r$ and $u, v = 1, \cdots, N$. Hence $\mathrm{Var}(\epsilon) = \Sigma \otimes \mathbf{I}_N$, an $Nr \times Nr$ matrix, where $\otimes$ means Kronecker product and $\mathbf{I}_N$ is an $N \times N$ identity matrix.

If we assume that the true or "feared" relationship over O, is represented as a quadratic model to each response, then

$$\mathrm{E}(\mathbf{Y}) \;=\; \eta \;=\; \mathbf{X}\beta \;+\; \mathbf{Z\Gamma} \;, \qquad (1.2)$$

where $\eta = (\eta_1, \cdots, \eta_r)$ with $\eta_i = (\eta_{1i}, \cdots, \eta_{Ni})'$, where $\mathbf{Z} = (\mathbf{z_1}', \cdots, \mathbf{z_N}')'$ with $\mathbf{z_u}' = (x_{u1}{}^2, \cdots, x_{uk}{}^2 \;:\; x_{u1}x_{u2}, \cdots, x_{uk}x_{u(k-1)})'$, and where $\Gamma = (\gamma_1, \cdots, \gamma_r)$

with $\gamma_i = (\beta_{11,i}, \beta_{22,i}, \cdots, \beta_{kk,i} \ : \beta_{12,i}, \cdots, \beta_{(k-1)k,i})'$ for $i = 1, \cdots, r$ and $u = 1, \cdots, N$.

It is expected that there would be a difference between the fitted model (1.1) and the feared function (1.2). Hence the criterion measuring the difference is needed and Box and Draper criterion can be extended to multiple response case. There is no restriction that all the experimental runs need to be within R in order to explore R.

2. CRITERION

2.1 The Form of Criterion

Let $\hat{\mathbf{y}}(\mathbf{x})$ and $\eta(\mathbf{x})$ be defined, respectively as fitted values of $\mathbf{Y}$ and the corresponding true mean values at the point $\mathbf{x}' = (1, x_1, \cdots, x_k)'$. In order to choose the design minimizing the errors, we use the following criterion measuring the amount of errors.

$$\mathbf{J} = \mathrm{N}\mathbf{\Sigma}^{-1} \int_{\mathrm{O}} \mathrm{w}(\mathbf{x}) \ \mathrm{E}\{\hat{\mathbf{y}}(\mathbf{x}) - \eta(\mathbf{x})\}'\{\hat{\mathbf{y}}(\mathbf{x}) - \eta(\mathbf{x})\} \ d\mathbf{x} \ , \tag{2.1}$$

where $w(\mathbf{x}) = \mathbf{\Omega} = \mathbf{1}/\int_{\mathbf{R}} \mathbf{dx}$ and 0 elsewhere, is a uniform weight function which gives weights to responses equally within R and $d\mathbf{x} = dx_1 \cdots dx_k$. The form of $\mathbf{J}$ can be written as following ;

$$\begin{aligned} \mathbf{J} &= \{ N\Omega \int_R \mathbf{x}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}' \ d\mathbf{x} \ \} \ \mathbf{I}_r \\ &\quad + \ N\Omega\mathbf{\Sigma}^{-1} \int_R \mathbf{\Gamma}'\{\mathbf{x}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z} - \mathbf{z}\}'\{\mathbf{x}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z} - \mathbf{z}\}\mathbf{\Gamma} \ d\mathbf{x} \\ &= \quad \mathbf{V} \quad + \quad \mathbf{B} \ , \end{aligned} \tag{2.2}$$

where $\mathbf{I}_r$ is an $r \times r$ identity matrix. From this, we see that the criterion $\mathbf{J}$, which is an $r \times r$ matrix, can be represented by the sum of the contributions from the two types of errors where $\mathbf{V}$ explains the variance error and $\mathbf{B}$ the bias error.

In minimizing $\mathbf{J}$, which is an $r \times r$ matrix , the trace of $\mathbf{J}$ can be used, say $tr(\mathbf{J})$. Using the trace of $\mathbf{J}$,

$$tr(\mathbf{J}) \ = \ tr(\mathbf{V} \ + \ \mathbf{B}) \ = \ tr(\mathbf{V}) \ + \ tr(\mathbf{B}) \ , \tag{2.3}$$

where $tr(\mathbf{V})$ and $tr(\mathbf{B})$ specify the amount of variance error and the bias error, respectively. Also assume that the design points are centrally located in the region R, i.e.,

$$\sum_{u=1}^{N} x_{ul} = 0 \text{ for } l = 1, \cdots, k. \tag{2.4}$$

Define $\{c^{lm}\} = \{c_{lm}\}^{-1} = \mathbf{C}^{-1} = N(\mathbf{X}'\mathbf{X})^{-1}$, where $c_{lm} = N^{-1}\sum_{u=1}^{N} x_{ul}x_{um}$ for $l, m = 0, 1, \cdots, k$ with $x_{ul} = 1$ for $l = 0$. Then

$$\begin{aligned} tr(\mathbf{V}) &= tr\{N\Omega \int_R \mathbf{x}(X'X)^{-1}\mathbf{x}'d\mathbf{x}\ \mathbf{I}_r\} \\ &= r\{1 + \sum_{l=1}^{k} c^{ll}/(k+2)\}. \end{aligned} \tag{2.5}$$

For second part of $tr(\mathbf{J})$, we define

$$\mathbf{\Sigma}^{-1} = \{\sigma^{ij}\} = |\mathbf{\Sigma}|^*\{\sigma^{ij}\sigma_i\sigma_j{}^{-1}\} \text{ for i, j} = 1, \cdots, \text{r}, \tag{2.6}$$

where

$$|\mathbf{\Sigma}|^* = \prod_{i=1}^{r} \sigma_i{}^2 |\mathbf{\Sigma}|^{-1} \tag{2.7}$$

with $|\mathbf{\Sigma}|$ as the determinant of $\mathbf{\Sigma}$. Also we define

$$\alpha_{lm,i} = \frac{\beta_{lm,i}}{\sigma_i/\sqrt{N}}, \tag{2.8}$$

which measures the ratio of the quadratic curvature term $\beta_{lm,i}$ to the sampling error for the mean of the i^{th} response, $\bar{y}_i$. Also define that

$$\delta_{1i} = \sum_{s=1}^{r} \sigma^{is} \left[\frac{1}{k+2}\sum_{g=1}^{k} \{\sum_{l=1}^{k}\sum_{m=l}^{k} \alpha_{lm,s}(*)\}\ \{\sum_{l=1}^{k}\sum_{m=l}^{k} \alpha_{lm,i}(*)\}\right] \tag{2.9}$$

with $(*) = \sum_{h=1}^{k} c^{gh}\ [hlm]$, where $[hlm] = N^{-1}\sum_{u=1}^{N} x_{uh}x_{ul}x_{um}$ is a third order design moment,

$$\delta_{2i} = \sum_{s=1}^{r} \sigma^{is}\ \{\sum_{l=1}^{k}\sum_{m=l}^{k} \alpha_{lm,s}\ (**)\}\ \{\sum_{l=1}^{k}\sum_{m=l}^{k} \alpha_{lm,i}\ (**)\} \tag{2.10}$$

with $(**) = c_{lm} - \delta_{lm}\ (k+2)$ and $\delta_{lm} = 1$ if $l = m$, 0, otherwise for $l, m = 1, \cdots, k$, and

$$\delta_{3i} = \sum_{s=1}^{r} \sigma^{is} \frac{(***)}{(k+2)^2(k+4)}, \tag{2.11}$$

with

$$(***) = 2(k+2)\sum_{l=1}^{k} \alpha_{lm,s}\alpha_{lm,i} + (k+2)\sum_{l=1}^{k}\sum_{m=l+1}^{k} \alpha_{lm,s}\alpha_{lm,i}$$
$$2(\sum_{l=1}^{k} \alpha_{ll,s})(\sum_{l=1}^{k} \alpha_{ll,i}). \tag{2.12}$$

Hence we can write

$$tr(\mathbf{B}) = |\mathbf{\Sigma}|^* \sum_{i=1}^{r} (\delta_{1i} + \delta_{2i} + \delta_{3i}). \tag{2.13}$$

Equation (2.4) can be written as following ;

$$\begin{aligned} tr(\mathbf{J}) &= tr(\mathbf{V}) + tr(\mathbf{B}) \\ &= r\{1 + \sum_{l=1}^{k} c^{ll}/(k+2)\} + |\mathbf{\Sigma}|^* \sum_{i=1}^{r} (\delta_{1i} + \delta_{2i} + \delta_{3i}) \end{aligned} \tag{2.14}$$

2.2 Minimizing the tr(J)

It is usually assumed that there exist the variance error and the bias error. Hence we consider the criterion which minimizes two types of errors simultaneously. In doing this, if we know the values of $\alpha_{lm,i}$'s, the optimum values of c_{lm} could be found, but usually α's are unknown. So we consider taking rotational average of $tr(\mathbf{J})$ over all orthogonal rotations of the response surface. By taking rotational average, $tr(\mathbf{V})$ is not changed, because it does not contain any α's but $tr(\mathbf{B})$ is.

The average values of $tr(\mathbf{B})$ over all rotations of the surface is obtained by substituting the average values of all rotations of the product $\alpha_{gh,i}\alpha_{lm,j}$. The detail for taking the rotational average are shown in Box and Draper (1959, Appendix 2).

With the third order moments zero, the form of $tr(\mathbf{J})$ we minimize is

$$\begin{aligned} tr(\mathbf{J}) &= r\left[1 + \frac{k}{(k+2)c}\right] \\ &+ \sum_{i,s=1}^{r} \sigma^{is} \left[(\sum_{l=1}^{k} \alpha_{ll,i})(\sum_{l=1}^{k} \alpha_{ll,s})\{c - 1/(k+2)\}^2 + (****)\right] \end{aligned}$$

$$= \quad rv(c) \; + \; \sum_{i,s=1}^{r} \sigma^{is} \left[(\sum^{k} \alpha_{ll,i})(\sum^{k} \alpha_{ll,s}) \; b(c) \; + (****) \right] \; , \quad (2.15)$$

where we can put $c_{ll} = c_{gg} = c$ for $l,\ g = 1, \cdots, k$ as a result of minimizing the rotational average of $tr(\mathbf{J})$ and c_{ij} is the second order design moment shown in (2.5), $b(c) = \{c - 1/(k+2)\}^2$ and

$$(****) \quad = \quad (k+2)^{-2}(k+4)^{-1} \left[2(k+2)\sum_{l=1}^{k} \alpha_{ll,i}\alpha_{ll,s} \right. \qquad (2.16)$$

$$\left. + \quad (k+2)\sum_{l<m}^{k} \alpha_{lm,i}\alpha_{lm,s} \; - \; 2(\sum_{l=1}^{k} \alpha_{ll,i})(\sum_{l=1}^{k} \alpha_{ll,s}) \right] . \quad (2.17)$$

The design we find in this case is a first order orthogonal design with the third moments zero and the second order design moments all equal, which is the same result as in the single response case.

We see that if the variance error is dominant, that is, the gap between the fitted model and the feared one is explained mainly by the sampling error, the design should be expanded to the experimental range. And when the bias error is dominant, that is, the error is caused mainly by the incorrect fitting of the model so that $tr(\mathbf{B})$ dominates $tr(\mathbf{V})$, the design size should be determined by the second order design moment which equals to $1/(k+2)$ regardless of the number of responses, as shown in equation (2.15). Also we can consider the usual situation where both types of errors are present. In that case, the optimum design size can be found by minimizing the criterion after computing each amount of errors so that the correlations among the responses will be taken into an account. Although the correlation coefficients among the responses are not shown directly in the form of a criterion, they are hidden in σ^{is}, and will effect the optimum design size.

Specific case with three responses and one predictor is examined to show how the optimum design is affected by to the changes of the correlations among the responses in the following part.

3. CASE WITH $r = 3$ AND $k = 1$

The form of criterion is

$$tr(\mathbf{J}) \; = \; 3v(c) \; + \; tr(\mathbf{B}) \; = \; 3\{1 \; + \; 1/(3c)\} \; + \; tr(\mathbf{B}) \; , \qquad (3.1)$$

where

$$tr(\mathbf{B}) \;=\; \{(c-1/3)^2 \;-\; 4/45\}\;|\Sigma|^*\;[@] \qquad (3.2)$$

with $|\Sigma|^* \;=\; \{1-\rho_{12}{}^2-\rho_{13}{}^2-\rho_{23}{}^2+2\rho_{12}\rho_{13}\rho_{23}\}^{-1}$, and where $\rho_{ij} \;=\; corr(y_i, y_j)$ for $i,\; j=1,\; 2,\; 3$, and

$$\begin{aligned}[@] \;=\;& (1-\rho_{23}{}^2)\alpha_{11,1}{}^2 \;+\; (1-\rho_{13}{}^2)\alpha_{11,2}{}^2 \;+\; (1-\rho_{12}{}^2)\alpha_{11,3}{}^2 \\ &+ 2\{(\rho_{13}\rho_{23}-\rho_{12})\alpha_{11,1}\alpha_{11,2} \;+\; (\rho_{12}\rho_{23}-\rho_{13})\alpha_{11,1}\alpha_{11,3} \\ &+ (\rho_{12}\rho_{13}-\rho_{23})\alpha_{11,2}\alpha_{11,3}\}.\end{aligned}$$

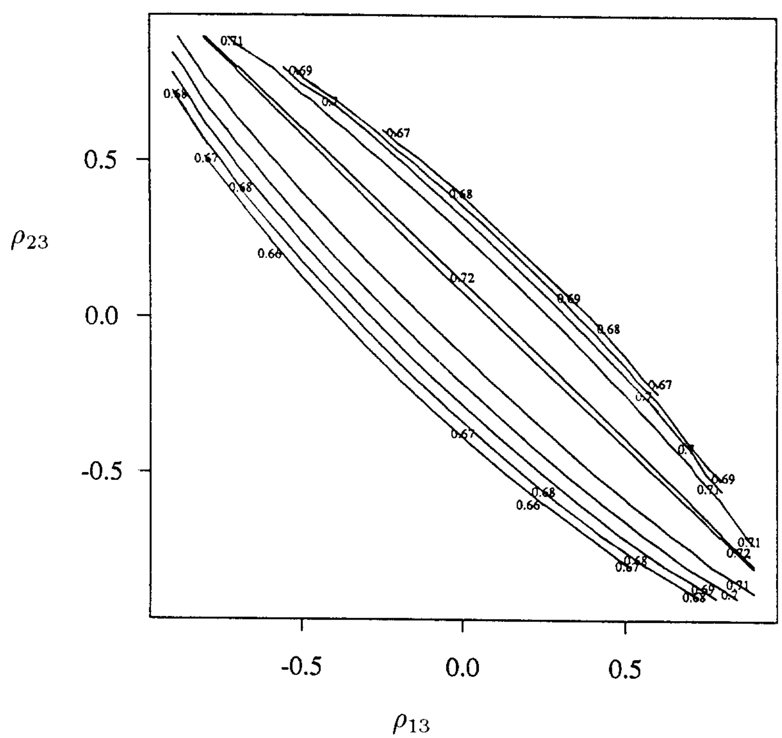

Figure 1. Change of optimum $\sqrt{c}$ with $\rho_{12} = -.9$

When the bias error is dominant, the optimum value of $\sqrt{c}$ is $\sqrt{1/3}$. Suppose that we consider the problem of choosing a design when there exists moderate amount of both types of errors. If we assume that $\alpha_{11,1} = \alpha_{11,2} = \alpha_{11,3} = 1$, then this could mean that the amount of the bias error and the variance one are almost same, from the definition of α's in (2.8). Then under

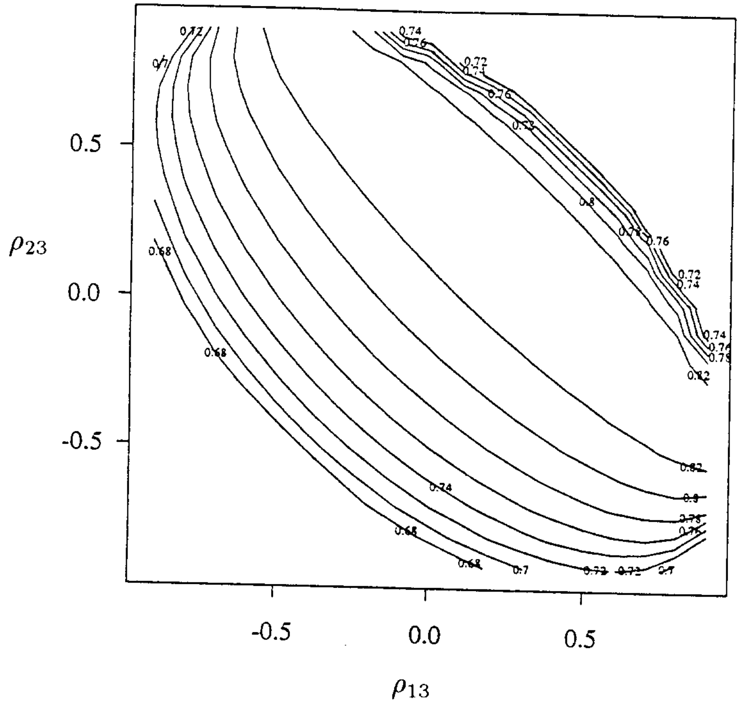

Figure 2. Change of optimum $\sqrt{c}$ with $\rho_{12} = -.5$

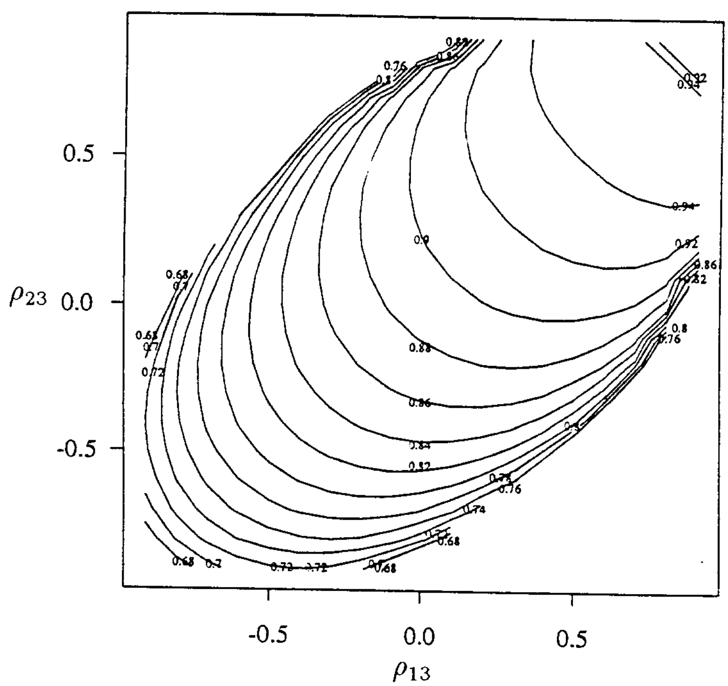

Figure 3. Change of optimum $\sqrt{c}$ with $\rho_{12} = .5$

these assumption, with $\rho_{12} = -.9,\ ,-.5,\ .5, \text{and},\ .9$, as ρ_{13} and ρ_{23} changes from -.9 to .9, FIGURES 1 through 4 are given to show change of optimum values of $\sqrt{c}$. The contours in Figures are drawn with possible range of ρ's which make $|\Sigma|$ positive From the plots, we find out the contours going down to the S-W direction($\rho_{13} < 0$ and $\rho_{23} > 0$) have smaller $\sqrt{c}$, that is, there is more bias present. When ρ_{12} has negative values(FIGURE 1 and FIGURE 2), there are changes in the N-E direction($\rho_{13} > 0$ and $\rho_{23} > 0$, where the $\sqrt{c}$ values are decreasing more steeply than they do in the S-W direction. When ρ_{13} and ρ_{23} have opposite signs, there is change in $\sqrt{c}$ in either the N-W direction or the S-E one. As ρ_{12} approaches 1, the bias values get smaller and the changes in $\sqrt{c}$ in the S-W direction get smoother, forming a downhill shape, and we have steep edge change in both S-E direction and the N-W direction. overall, we see that the more negative the correlation, the more bias is obtained when the three α's are the same, and the smaller are the designs.

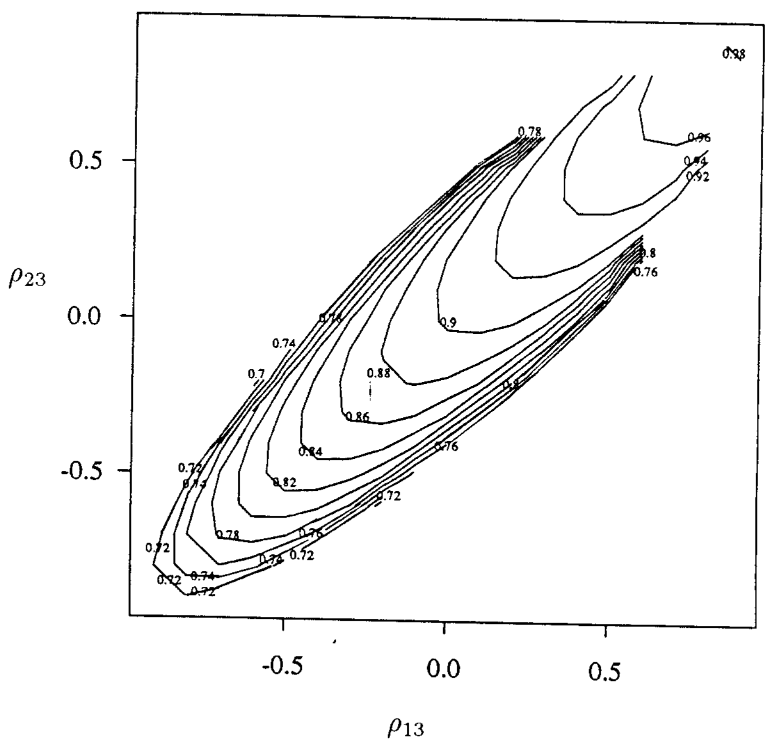

Figure 4. Change of optimum $\sqrt{c}$ with $\rho_{12} = .9$

4. CONCLUDING REMARKS

The optimum design was chosen by minimizing the matrix $\mathbf{J} = \mathbf{V} + \mathbf{B}$, using the $tr(\mathbf{J})$. The optimum design is a first order design with the third order design moments zero and the second order moments all equal, so that the suitable design size is determined by this common second order design moment. With dominant bias, the optimum values of the second order design moment is $1/(k+2)$ as in the single response case. With proper amount of bias and variance error, the design is affected by the correlations more or less as shown in FIGURES 1 through 4.

REFERENCES

(1) Box, G.E.P. and Draper, N.R. (1959). A basis for the selection of a response surface design, *Journal of American Statistical Associaion*, **54**, 622-654.

(2) Box, G.E.P. and Draper, N.R. (1963). The choice of a second order rotatable designs, *Biometrika*, **50**, 335-352.

(3) Box, G.E.P. and Tiao, G.C. (1973). *Bayesian Inference in Statistical Analysis.*, Wiley, New york.

(4) Khuri, A.I. (1988). The analysis of multiresponse experiments: a review, *Technical Report 302* , Dept. of Stat., Univ. of Florida, Gainesville.

(5) Kim, W.B. and Draper, N.R. (1994). Choosing a design for straight line fits to two correlated responses, *Statistica Sinica*, **4**(1), 275-280.